

Supplement: Iterative Selection Regression (Maller *et al.*, 1983)

Procedure fits a lower bound to a data set. The procedure is essentially the same to fit an upper limit line. This Supplement relates to the use of the Excel spreadsheet 'Maller Iterative Selective Regression'.

Step 1

Input **lower points** of the dataset as (x_i, y_i) ; Carry out **linear regression** of (x_i, y_i) , shown as in the excel figure; Calculate the **simulated y value** according to the linear regression; Calculate **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i - x_{ave})^2$** ; When **$y_i < y$** , calculate **$y_i(x_i - x_{ave})$** and the sum value, when **$y_i > y$** , calculate **$y(x_i - x_{ave})$** and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i - x_{ave}) + \text{sum of } y(x_i - x_{ave})) / \text{sum of } (x_i - x_{ave})^2$ "; According to the point which has the highest $(y_i - y)$, calculate the **intercept of the new fitting line**. In this step, we get the new fitting line " $y = 0.439542023x + 10.05424883$ " for use in next step. Besides, the y_i which has the highest $(y_i - y)$ value should be changed as y in next step.

Step 2

The y_i which has the highest $(y_i - y)$ in step 1 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 1 " $y = 0.439542023x + 10.05424883$ "; Calculate **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i - x_{ave})^2$** ; When **$y_i < y$** , calculate **$y_i(x_i - x_{ave})$** and the sum value, when **$y_i > y$** , calculate **$y(x_i - x_{ave})$** and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i - x_{ave}) + \text{sum of } y(x_i - x_{ave})) / \text{sum of } (x_i - x_{ave})^2$ "; According to the point which has the highest $(y_i - y)$, calculate the **intercept of the new fitting line**. In this step, we get the new fitting line " $y = 0.396965376x + 11.08800981$ " for use in next step. Besides, the y_i which has the highest $(y_i - y)$ value should be changed as y in next step.

Step 3

The y_i which has the highest $(y_i - y)$ in step 2 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 2 " $y = 0.396965376x + 11.08800981$ "; Calculate the **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i - x_{ave})^2$** ; When $y_i < y$, calculate $y_i(x_i - x_{ave})$ and the sum value, when $y_i > y$, calculate $y(x_i - x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i - x_{ave}) + \text{sum of } y(x_i - x_{ave})) / \text{sum of } (x_i - x_{ave})^2$ "; According to the point which has the highest $(y_i - y)$, calculate the **intercept of the new fitting line**. In this step, we get the new fitting line " $y = 0.361582237x + 18.29484759$ " for use in next step. Besides, the y_i which has the highest $(y_i - y)$ value should be changed as y in next step.

Step 4

The y_i which has the highest $(y_i - y)$ in step 3 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 3 " $y = 0.361582237x + 18.29484759$ "; Calculate the **the difference between y_i and y** ; As all $y_i < y$, Carry out **linear regression** of (x_i, y_i) , shown as the figure. In this step, we get the new fitting line " $y = 0.3606x + 8.8597$ " for use in next step.

Step 5

Calculate **simulated y value** according to linear regression in step 4 " $y = 0.3606x + 8.8597$ "; Calculate the **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i - x_{ave})^2$** ; When $y_i < y$, calculate $y_i(x_i - x_{ave})$ and the sum value, when $y_i > y$, calculate $y(x_i - x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i - x_{ave}) + \text{sum of } y(x_i - x_{ave})) / \text{sum of } (x_i - x_{ave})^2$ "; According to the point which has the highest $(y_i - y)$, calculate the **intercept of the new fitting line**.

In this step, we get the new fitting line " $y = 0.335008828x + 14.07210984$ " for use in next step. Besides, the y_i which has the highest $(y_i - y)$ value should be changed as y in next step.

Step 6

The y_i which has the highest $(y_i - y)$ in step 5 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 5 " $y = 0.335008828x + 14.07210984$ "; Calculate the **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i - x_{ave})^2$** ; When $y_i < y$, calculate $y_i(x_i - x_{ave})$ and the sum value, when $y_i > y$, calculate $y(x_i - x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i - x_{ave}) + \text{sum of } y(x_i - x_{ave})) / \text{sum of } (x_i - x_{ave})^2$ "; According to the point which has the highest $(y_i - y)$, calculate the **intercept of the new fitting line**. In this step, we get the new fitting line " $y = 0.317109702x + 15.25577904$ " for use in next step. Besides, the y_i which has the highest $(y_i - y)$ should be changed as y in next step.

Step 7

The y_i which has the highest $(y_i - y)$ in step 6 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 6 " $y = 0.317109702x + 15.25577904$ "; Calculate the **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i - x_{ave})^2$** ; When $y_i < y$, calculate $y_i(x_i - x_{ave})$ and the sum value, when $y_i > y$, calculate $y(x_i - x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i - x_{ave}) + \text{sum of } y(x_i - x_{ave})) / \text{sum of } (x_i - x_{ave})^2$ "; According to the point which has the highest $(y_i - y)$, calculate the **intercept of the new fitting line**. In this step, we get the new fitting line " $y = 0.305387181x + 17.64342214$ " for use in next step. Besides, the y_i which has the highest $(y_i - y)$ should be changed as y in next step.

Step 8

The y_i which has the highest $(y_i - y)$ in step 7 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 7 " $y = 0.305387181x + 17.64342214$ "; Calculate the **the difference between y_i and y** ; As all $y_i < y$, Carry out **linear regression** of (x_i, y_i) , shown as in the figure. In this step, we get the new fitting line " $y = 0.3054x + 12.352$ " for use in next step.

Step 9

Calculate **simulated y value** according to linear regression in step 8 " $y=0.3054*x+12.352$ "; Calculate the **difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i-x_{ave})^2$** ; When $y_i < y$, calculate $y_i*(x_i-x_{ave})$ and the sum value, when $y_i > y$, calculate $y*(x_i-x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i*(x_i-x_{ave}) + \text{sum of } y*(x_i-x_{ave}))/\text{sum of } (x_i-x_{ave})^2$ "; According to the point which has the highest (y_i-y) , calculate the **intercept of the new fitting line**.

In this step, we get the new fitting line " $y=0.289271609*x+15.6372306$ " for use in next step. Besides, the y_i which has the highest (y_i-y) should be changed as y in next step.

Step 10

The y_i which has the highest (y_i-y) in step 9 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 9 " $y=0.289271609*x+15.6372306$ "; Calculate the **difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i-x_{ave})^2$** ; When $y_i < y$, calculate $y_i*(x_i-x_{ave})$ and the sum value, when $y_i > y$, calculate $y*(x_i-x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i*(x_i-x_{ave}) + \text{sum of } y*(x_i-x_{ave}))/\text{sum of } (x_i-x_{ave})^2$ "; According to the point which has the highest (y_i-y) , calculate the **intercept of the new fitting line**. In this step, we get the new fitting line " $y=0.282076325*x_i+16.11305474$ " for use in next step. Besides, the y_i which has the highest (y_i-y) should be changed as y in next step.

Step 11

The y_i which has the highest (y_i-y) in step 10 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 10 " $y=0.282076325*x_i+16.11305474$ "; Calculate the **difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i-x_{ave})^2$** ; When $y_i < y$, calculate $y_i*(x_i-x_{ave})$ and the sum value, when $y_i > y$, calculate $y*(x_i-x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i*(x_i-x_{ave}) + \text{sum of } y*(x_i-x_{ave}))/\text{sum of } (x_i-x_{ave})^2$ "; According to the point which has the highest (y_i-y) , calculate the **intercept of the new fitting line**.

In this step, we get the new fitting line " $y=0.361582237x+18.29484759$ " for use in next step. Besides, the y_i which has the highest (y_i-y) value should be changed as y in next step.

Step 12

The y_i which has the highest (y_i-y) in step 11 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 11 " $y=0.361582237x+18.29484759$ "; Calculate **the difference between y_i and y** ; Calculate the **average value of x_i** ; Calculate the **difference between x_i and x_{ave}** ; Calculate the **sum of $(x_i-x_{ave})^2$** ; When $y_i < y$, calculate $y_i(x_i-x_{ave})$ and the sum value, when $y_i > y$, calculate $y(x_i-x_{ave})$ and the sum value; Calculate the **slope of new fitting line** according to " $(\text{sum of } y_i(x_i-x_{ave}) + \text{sum of } y(x_i-x_{ave})) / \text{sum of } (x_i-x_{ave})^2$ "; According to the point which has the highest (y_i-y) , calculate the **intercept of the new fitting line**.

In this step, we get the new fitting line " $y=0.361582237x+18.29484759$ " for use in next step. Besides, the y_i which has the highest (y_i-y) value should be changed as y in next step.

Step 13

The y_i which has the highest (y_i-y) in step 12 should be changed as y ; Calculate the **simulated y value** according to the calculated fitting line in step 12 " $y=0.361582237x+18.29484759$ "; Calculate **the difference between y_i and y** ; As all $y_i < y$, Carry out **linear regression** of (x_i, y_i) , shown as the figure. In this step, we get the new fitting line " $y=0.2774x+13.998$ ".

As $R^2=0.9577 > 0.95$, we get the final lower boundary " $y=0.2774x+13.998$ ".